

CRF Errors Corrected by the STIC Systems Branch

Serial Number: 09/840,746

CRF Processing Date: 5/10/2001
Edited by: *[Signature]*
Verified by: *[Signature]* (STIC staff) Changed a file from non-ASCII to ASCII**ENTERED** Changed the margins in cases where the sequence text was "wrapped" down to the next line. Edited a format error in the Current Application Data section, specifically: Edited the Current Application Data section with the actual current number. The number inputted by the applicant was the prior application data; or other _____. Added the mandatory heading and subheadings for "Current Application Data". Edited the "Number of Sequences" field. The applicant spelled out a number instead of using an integer. Changed the spelling of a mandatory field (the headings or subheadings), specifically: Corrected the SEQ ID NO when obviously incorrect. The sequence numbers that were edited were: Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited: Corrected subheading placement. All responses must be on the same line as each subheading. If the applicant placed a response below the subheading, this was moved to its appropriate place. Inserted colons after headings/subheadings. Headings edited included: Deleted extra, invalid, headings used by an applicant, specifically: Deleted: non-ASCII "garbage" at the beginning/end of files; secretary initials/filename at end of file; page numbers throughout text; other invalid text, such as _____. Inserted mandatory headings, specifically: Corrected an obvious error in the response, specifically: Edited identifiers where upper case is used but lower case is required, or vice versa. Corrected an error in the Number of Sequences field, specifically: A "Hard Page Break" code was inserted by the applicant. All occurrences had to be deleted. Deleted *ending* stop codon in amino acid sequences and adjusted the "(A)Length:" field accordingly (error due to a PatentIn bug). Sequences corrected: Other:

OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/840,746

DATE: 05/10/2001
TIME: 09:48:44

Input Set : A:\pc0039us_seqlist.txt
Output Set: N:\CRF3\05102001\I840746.raw

Does Not Comply
Corrected Diskette Needed

2 <110> APPLICANT: Chen, Huei-Mei
3 Honchell, Cynthia D.
4 Tang, Y. Tom
6 <120> TITLE OF INVENTION: Mucin-Related Tumor Marker
8 <130> FILE REFERENCE: PC-0039 US
C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/840,746
C--> 11 <141> CURRENT FILING DATE: 2001-04-23
13 <160> NUMBER OF SEQ ID NOS: 20
14 <170> SOFTWARE: PERL Program

ERRORED SEQUENCES

155 <210> SEQ ID NO: 2
156 <211> LENGTH: 6952
157 <212> TYPE: DNA
158 <213> ORGANISM: Homo sapiens
160 <220> FEATURE:
161 <221> NAME/KEY: misc_feature
162 <223> OTHER INFORMATION: Incyte ID No: 182514CB1
164 <400> SEQUENCE: 2
165 gttcgatgaa agaattgccg ctttcaaacc aaagagtggaa acagcctcgagatggaaac 60
166 agagagggcg atggggctgt cagaagaatg gactgtgcac agccaagagg ccaccacttc 120
167 ggcttggagc cttcccttcc ttctctgttt ggagatggaa gagctgacca cgccttctag 180
168 gaagagaaat tcctcaggac cagatctctc ctggctgcat ttctacagga cagcagcttc 240
169 ctctcctctc tttagaccctt cctcaccttc taaaagtata gagaagctta acaactccac 300
170 tggcctccag agctcctcag tcagtccaaac aaagacaatg catgttgcata cctgttccac 360
171 ttagtggcgc ccgagaacgc tgcatcttt gacggtcagt ctggacactg tgaccaagac 420
172 agaaggcttc cccaaaggact ccagaattgc cacgacttca tcctcagtc ttcttcacc 480
173 ctctcagtg gaatcggaaa gaaacagtag agtaactggg aatccagggg atgaggaatt 540
174 cattgaacca tccacagaaa atgaaattgg acttaacgtt ttgcgtggca aatgattcc 600
175 ccaacccatgg gagaacatca gcttgcacgc agctctgggg tgcaaaatgg aagtccatgg 660
176 tctcagactg agactgttc tagtcagtc gcacccatga gaggtggaga gatcactgca 720
177 cactggctct tgaccaacag cacaacatctc gcatgttgcata caggaagctc tgcttcata 780
178 cctgaagggt tgaatgttc agtggacc cagttctcag actctactgt acagtctgg 840
179 ggaagtca cagcattggg agataggat tatttcacatctac atcttcctcg 900
180 gaaagctca attcatcagc accacgtggaa gaaacgttca tcgtggat tagctacgg 960
E--> 181 caagtgcgtg gcacagctat tgaacaaagg acttccacgc accacacaga ccacacctac 1020
182 1020 *format error*
E--> 183 ctgtcatctca ctttccaaaggaggaaacgg gcgttactgt ccattacaga taacagttca
184 1080
E--> 185 tcctcagaca ttgtggagag ctcaacttct tatattaaaa tctcaaactc ttccatca
186 1140
E--> 187 gagtatttcctca tgctcagact gagagaagta acatctcata ctatgacggg
188 1200
E--> 189 gaatatgctc agccttctac tgagtcgcca gttctgcata catccaaacct tccgtcctac
190 1260

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/840,746

DATE: 05/10/2001

TIME: 09:48:44

Input Set : A:\pc0039us_seqlist.txt
 Output Set: N:\CRF3\05102001\I840746.raw

E--> 191 acacccacca ttaatatgcc gaacacttcg gttgttctgg acactgatgc tgagtttgtt
 192 1320
 E--> 193 agtgactcct cctcctcctc ttccctcctcc tcctcttctt cttcttcagg gcctcctttg
 194 1380
 E--> 195 cctctgcctc ctgtgtcaca atcccaccaat ttatttcat caatttacc atcaaccagg
 196 1440
 E--> 197 gcctctgtgc atctactaaa gtctacctct gatgcatcca caccatggc ttcctcacca
 198 1500
 E--> 199 tcacctttac cagtatcctt aacgacatct acatctgcc cactttctgt ctcacaaaca
 200 1560
 E--> 201 accttgcac agtcatcttc taccctgtc ctgcccaggg caagggagac tcctgtgact
 202 1620
 E--> 203 tcatccaga catcaacaat gacatcattc atgacaatgc tccatagtag tcaaactgca
 204 1680
 E--> 205 gacctaaga gccagagcac cccacaccaa gagaaagtca ttacagaatc aaagtcacca
 206 1740
 E--> 207 agcctgggtg ctctgcccac agagtccacc aaagctgtaa caacaaactc tcctttgcct
 208 1800
 E--> 209 ccattcctaa cagagtccctc cacagagcaa acccttccag ccacaagcac caacttagca
 210 1860
 E--> 211 caaatgtctc caactttcac aactaccatt ctgaagacct ctcagcctct tatgaccact
 212 1920
 E--> 213 cctggcaccc tgtcaagcac agcatctctg gtcactggcc ctatagccgt acagactaca
 214 1980
 E--> 215 gctggaaaac agctctcgct gacccatctt gaaatactag ttccctcaa at ctcacacagaa
 216 2040
 E--> 217 ggtggcatca gcacagaaag gaaccgagtg attgtggatg ctaccactgg attgatccct
 218 2100
 E--> 219 ttgaccagtg tacccacatc agcaaaagaa atgaccacaa agcttggcgt tacagcagag
 220 2160
 E--> 221 tacagccccag cttcacgttc cctcggaaaca tctccttctc cccaaaccac agttgtttcc
 222 2220
 E--> 223 acggctgaag acttggctcc caaatctgcc acctttgctg ttcagagcag cacacagtca
 224 2280
 E--> 225 ccaacaacac tgtcctcttc agcctcagtc aacagctgtg ctgtgaaccc ttgtcttcac
 226 2340
 E--> 227 aatggcgaat gctgcgaga caacaccagc cgtggctacc actgcagggtg cccgccttcc
 228 2400
 E--> 229 tggcaagggg atgattgcag tgtggatgtg aatgagtgcc tgtcgaaccc ctgcccattcc
 230 2460
 E--> 231 acagccacgt gcaacaatac tcagggatcc tttatctgca aatgcccgt tgggtaccag
 232 2520
 E--> 233 ttggaaaaag ggatatgcaa tttggatgaa accttcgtga cagagttaa attaaagaga
 234 2580
 E--> 235 actttctta atacaactgt ggaaaaacat tcagacctac aagaagttga aaatgagatc
 236 2640
 E--> 237 accaaaaacgt taaaatgtg ttttcagcg ttacctagtt acatccgatc tacagttcac
 238 2700
 E--> 239 gcctcttaggg agtccaaacgc ggtggatgc tcactgcaaa caacctttc cctggcctcc

Manne

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/840,746

DATE: 05/10/2001

TIME: 09:48:44

Input Set : A:\pc0039us_seqlist.txt
 Output Set: N:\CRF3\05102001\I840746.raw

240 2760
 E--> 241 aatgtgacgc tatttgacct ggctgatagg atgcagaaaat gtgtcaactc ctgcaagtcc
 242 2820
 E--> 243 tctgctgagg tctgccagct cttggatct cagaggcgg a tcttagagc gggcagctt
 244 2880
 E--> 245 tgcaagcgg a agagtcccga atgtgacaaa gacacctcca tctgactga cctggacggc
 246 2940
 E--> 247 gttgccctgt gccagtgcaa gtcgggatac tttcagttca acaagatgg a cactcctgc
 248 3000
 E--> 249 cgagcatgtg a agatggata taggcttcaa aatgaaacct gcatgagttt cccatggc
 250 3060
 E--> 251 cttgggtggc tcaactgtgg aaacccctat cagttatca ctgtgggtat cgcagccgc
 252 3120
 E--> 253 ggaggtgggc tcctgctcat cctaggcatc gcactgattt ttacctgtt cagaaagaat
 254 3180
 E--> 255 aaaaatgaca taagcaaact catcttcaaa agtggagatt tccaaatgtc cccatatgct
 256 3240
 E--> 257 gaatacccca aaaatcctcg ctcacaagaa tggggccgag aagctattga aatgcatgag
 258 3300
 E--> 259 aatggaaatgtc cccaaacccct cctccagatg acggatgtgt actactcgcc tacaagtgt
 260 3360
 E--> 261 aggaatccag aacttgaacg aaacggactc taccggcct acactggact gccaggatca
 262 3420
 E--> 263 cggcatttccc gcattttccc cggacagtat aaccctgtt tcatcagtga tgaaagcaga
 264 3480
 E--> 265 agaagagact acttttaagt ccaggagaga gagggactca ttgctctgag ccagtcaccc
 266 3540
 E--> 267 gggacctctg ctcagaggac cgcaccagga ggctgcgcaggatgtc gggagccacg
 268 3600
 E--> 269 ctgagtggca agcaggaaga gggacaggca tgcggggcgt gaccacagt gaggagacag
 270 3660
 E--> 271 gtggatgtgg aaccacaggc tgctcattca gcacctttgt tttactgtg aacgtgaatg
 272 3720
 E--> 273 tggccagta tcaagagagt ctctctgagt gactgcacca tggcaactggc accagggcga
 274 3780
 E--> 275 ctattagcca gggcagacca ctagacttca gtgcaggac ctggtttcc ctgcgttgc
 276 3840
 E--> 277 acttttagtaa attgggtggg aggtttcctt ttggatctgt tttgagactg ttccagaaaag
 278 3900
 E--> 279 aaggcttcct ttcccgagac acttccatag gcagcaattt ggtgattcat ttgcagcaaa
 280 3960
 E--> 281 atactggctt gttaattatt ttctgcccgcgcgt gctaaacaac agatgaggat
 282 4020
 E--> 283 gagcgtacca ctgaagtctg aagatgtcgc cattgaacgg acagtgtttt catatgttcc
 284 4080
 E--> 285 taggttgtct tatgtacag ttccaaagcc agccccacaca gtgaggaaat gtgtgaggca
 286 4140
 E--> 287 ccgcacacaa ctgcaatgtg ttttttaagt caaggtgaca catgtattta agatttttt
 288 4200

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/840,746

DATE: 05/10/2001
TIME: 09:48:44

Input Set : A:\pc0039us_seqlist.txt
Output Set: N:\CRF3\05102001\I840746.raw

E--> 289 ttaaaatctc tttgcagtt aatctcaactt tttcaaacaa gcctggatca gggcaaaaca
290 4260
E--> 291 acttatattt ggtttagct ggaggctcag caggcagatt gcaggcaggg gggcacttt
292 4320
E--> 293 catccatgag ggcccagcct ggggcctggg actctgatca ccattgtgga gccagaggc
294 4380
E--> 295 agctgcgtat ggaggagaaa tgtcaaactg aacgcaggtt tcaccactct aggaaagcag
296 4440
E--> 297 cttgttgagc ccctgcagct ggatgtgggt agagggatgg gctgaatagg caggttagat
298 4500
E--> 299 ttcctgcata aacagtgcct tggaaagctg tgtggattcc tgaggaagaa cagggagccg
300 4560
E--> 301 agatggagcc acacatgagt ttgctcaccg gctactgcag cactttgtac ccagaatctc
302 4620
E--> 303 atgtccacaa accccatgta aactttcaac cactcaaagc tgtttattcg gctgaagaaa
304 4680
E--> 305 taactttttt ttctcaccca gtcatttgta cctcttcata tggctgtgtc gcaccctcca
306 4740
E--> 307 gaaacgtggt tatacttcca gtcagtgtgg gagaactgaa gacttccggg tggtcgagga
308 4800
E--> 309 actgaggggtt gaccttcggg aaggaagttc cactcatctt atttattatg cctgtgtatgt
310 4860
E--> 311 gggcctgccc agggagacat ccagtactcg gtgtctttaa ttgccacctg gggactgtg
312 4920
E--> 313 ttatttggcc ttctttgggg catcctgggtt ttggatgaag tgagggaat acagaggtaa
314 4980
E--> 315 aagaattgtc tccaccctga agcggggagt cccgcttcac atttctggaa atgggtgcagc
316 5040
E--> 317 cactggggac agttctgccc cgggcatggt tgtttctca aggtcctcta aatataatcc
318 5100
E--> 319 ctatttcttac ataatccttg gccctgatgg ttttaagcaa gaactcctgt gtcccatgg
320 5160
E--> 321 ctccaccact caccatcacc ctgctgttagc aagagtccctaa gtcagggag gtgcatttt
322 5220
E--> 323 gtagttaaat tgcacttatac catgagataa ataaaaggag aactgtttttt atcagtggag
324 5280
E--> 325 gctaaccctaa aatttcaaag tgtcgccctt ttgaaatctt gggcctctct ctctgttagaa
326 5340
E--> 327 ccaatggccc tttgtggctc acggcctcgc acctaactgg agagttctga gtcctgcag
328 5400
E--> 329 ctcacctgag cccacagact aggcttcttg gtccttccg cagcatgcct gtcacccccc
330 5460
E--> 331 agaaccgcga gctgtggaa gagccatgta gggaggctat tcccaggcat acacttccac
332 5520
E--> 333 tgccttcagc tgacgtcaca gctgacaaat catctccct atcggagccaa gaagacttca
334 5580
E--> 335 gctccacaaa atgaagtgtt ctgtcctgaa aacattctt ggaagaatcc caacatcgag
336 5640
E--> 337 aaaacgggtt cctgtgagtt ccaacaatgc ttcttgcata tgggttctt ccgtatggag

Done

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/840,746

DATE: 05/10/2001

TIME: 09:48:44

Input Set : A:\pc0039us_seqlist.txt
Output Set: N:\CRF3\05102001\I840746.raw

338 5700
E--> 339 tggattaaga gtgtttatt ttgttgttct aactgagaaa aaaaggaggc acccacaagg
340 5760
E--> 341 ttgaggtcac acagtctcca cagttccag gaggcgttg ggggtgggga aggcacctcc
342 5820
E--> 343 agagcatgag gctctaaggg gacatgagta aagcatgtct gtgacccagt gaggaaggga
344 5880
E--> 345 gaggccagct gcactcctgc acggggttcc tagctgcaga agggtcccgc ctaggcccag
346 5940
E--> 347 gggaaacacc tgatagcaga agaggcctgg atgcacacct ggcacgcccga ggctctccgc
348 6000
E--> 349 ccagacacag tgctccatgt cagccctgc acctgggtg tgtgattcac gtgcacagat
350 6060
E--> 351 gccacaatcc tgcaccaata tcccacagat ggggaaaggt gagaggaagg ggcaagtgtat
352 6120
E--> 353 gtgttaactgc tcaagagatg cttaaacctc catagagagg agccggcgc aggggcattct
354 6180
E--> 355 gtgtgtcccg tcacacactg cagcagggaa ggggtggctgg ctggctccct ggcacatcgt
356 6240
E--> 357 gtttggttta agctccagag ggtcttattt ccattgtctt ttcccttgcc ctttgagcca
358 6300
E--> 359 gcctaaggcc ctggagtcgt tttctttagg cggatgaact gacatgtcc taccatgacc
360 6360
E--> 361 aggctctggg caaggctctt cacagtatcc ttgagaggtg ggcacatggaa tgcccatattc
362 6420
E--> 363 tcaggtacag aaaccttcag agaggataaa tagcttgcctc tgtagaagca ggactgaaac
364 6480
E--> 365 ctttgcgtccgc ctgactcccc cagctactct gcccactgtt gccccctgccc ttactgtcct
366 6540
E--> 367 ggcacacccc tcaccatctt gtataccttta aatatcaaag agggcaagag agaaaggc
368 6600
E--> 369 ttaaagataaa gttatttttt taaggaacct taatattatt ttaagaagt aaccaaatta
370 6660
E--> 371 gtgacgtgaa atgcaaaaaa aaaaaaaaaa aatgtctact accctttgtt aatgtgtt
372 6720
E--> 373 tcagattgtt ttttatatgt aattcttaga cacttgtcat taagaaaata gtggctggct
374 6780
E--> 375 tttgtgttcagc aagaaggcaca ctggcacgtg gctttggat aggaagtggaa aggcaaggac
376 6840
E--> 377 ctgggtttct gacaagtgcc gtcagactta cccttccatc tggagagctg gtggctttgg
378 6900
E--> 379 tcccctgggt agggccatgg gttccccact attactggaa agctataaggg tg
380 6952

Same

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/840,746

DATE: 05/10/2001
TIME: 09:48:45

Input Set : A:\pc0039us_seqlist.txt
Output Set: N:\CRF3\05102001\I840746.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application Number
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:181 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:1020 SEQ:2
M:254 Repeated in SeqNo=2
L:658 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15
L:659 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15
L:678 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16
L:679 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16
L:680 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16
L:681 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16
L:700 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17
L:702 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17